

DRAFT - Langley DAAC User Working Group (UWG)

Agenda

Preliminary Agenda

Location: DAAC Conference Room - Building 1268C, Room 1313

Date: October 17 - 18, 2001

Please Note:

- Times are approximate; breaks will take place but are not listed.
- Security into NASA Langley is significantly enhanced. Be sure to leave extra time to obtain visitor badges.

Wednesday, October 17, 2001

Time	Item	Presenter(s)	Description
9:00 a.m.	Introduction/Welcome	J. S. Levine / C. Randall	Introduction of new UWG members: Mike Daniels, NCAR, Kyle Miller, JPL, and Glenn Schuster, U. S. Satellite Laboratory
9:15 a.m.	Change and Continuity	B. R. Barkstrom	Management changes at the ASDC and how the organization has dealt with continuity during the last few months.
10:00 a.m.	ASDC Work Plan Discussion	J. S. Levine / C. Randall, Moderators	
11:15 a.m.	NOAA Long Term Archival (LTA) Plans	B. R. Barkstrom	The large scale political situation, immediate activities, the current version of the Plan, possible actions
11:45 a.m.	LTA Plan Discussion	J. S. Levine / C. Randall, Moderators	Key point: UWG position on LTA Plan
12:30 p.m.	Lunch		
1:30 p.m.	ERBE Archival Activities	M. T. Ferebee	Budget process reduced V0, but GSFC has wanted us to take on moving ERBE from GSFC - we also need to deal with how to archive ERBE printed material and ground cal results
2:00 p.m.	New Vision and New Activities	B. R. Barkstrom	A discussion of how our vision of the ASDC needs to respond to changing conditions, including the end of the ECS contract, current state of the ESIP Federation, the EMD activity, NewDISS, and changes in the IT world.
3:00 p.m.	Special Presentation	Greg Carmichael	The NASA TRACE-P Mission by Greg Carmichael, Professor, Department of Chemical Engineering, Center for Global and Regional Environmental Research, University of Iowa.
4:00 p.m.	Using Web-DAV for Collaborative Web	J. Olson	Authoring - experience with a new Tool



Thursday, October 18, 2001

Time	Item	Presenter(s)	Description
9:00 a.m.	UWG Activities, including the User Survey	J. Olson or other	
10:00 a.m.	Old Business/New Business	C. Randfall / J. S. Levine	
11:30 a.m.	Special Lecture	Bruce Wielicki	The Latest Results from CERES by Bruce Wielicki, NASA Langley (Pearl Young Theater)
12:30 p.m.	Lunch		
1:30 p.m.	Closed Session of UWG	(UWG Members Only)	
2:30 p.m.	Report back to DAAC Management		
3:00 p.m.	Adjourn		

Some Comments on the Long Term Archival of Data from Bruce Barkstrom:

Here's a suggestion of the issues that seem most important about the Long Term Archive Plan from my perspective. I will try to get a copy of that Plan to the UWG members before the meeting and see if they would be willing to endorse something like the following statement. I'll also try to get you a copy. The Plan does suggest that NASA would agree to start transferring EOS instrument data to NOAA for Long Term in about 2006, with complete transfer by 2012. I think this is a matter of concern with respect to the long term viability of the ASDC.

Bruce B.

1. It is critical to ensure that the Long Term Archival Plan have realistic estimates of the amount of computation and manpower resources needed for systematic reprocessing of data to ensure data product improvement to climate quality.

While the draft LTA Plan incorporates archival of Level 0 data, scientific work is most likely to proceed using Level 2 and Level 3 data. Accordingly, provision must be made to include these higher level data in any LTA Plan. The EOS instrument teams have already invested large amounts of time and other resources into producing high quality data. The production experience of these teams suggests that higher level data production is complex and requires very careful attention to rigorous configuration management and quality assurance. There is a concern that the LTA job may be viewed as one in which the least expensive approach to providing higher level data would be to start with the Level 0 data and the science team software. Our experience suggests that this approach may well require an investment approaching the cost of the EOS science team software development and of the operations expense of rerunning the EOS data production. The experience of both the Instrument Teams and the Data Production centers must be incorporated into the cost estimates for realistic scenarios for Long Term archival and for future reprocessing.

2. It is critical to the long term validity of the data that reprocessing or reproduction efforts have realistic estimates of the cost of revalidation and quality assurance.

Long Term Archival of scientific data will require either preserving the original data products or methods of



revalidating the data after reprocessing. Thus, responsible data stewardship requires including the cost of this revalidation effort in the cost of the LTA job. Adequate revalidation may require archival of additional ancillary data that were used to provide vicarious instrument calibrations or other "ground truth". It also seems likely that the LTA job will also include the archival of ground calibration data and other kinds of records that provide traceability to fundamental physical standards. In some cases, revalidation may also require large-scale consistency checks of higher level data products. The LTA job needs to incorporate systematic planning for both archival of the appropriate ancillary data and for the revalidation activity.

3. The Long Term Archival Plan needs to ensure that there is referential integrity of the data.

Scientific discoveries from the EOS data and other data in the Long Term Archive need to cite the original data from which conclusions are drawn. The citations cannot change as a result of migration to a Long Term Archive. Not only does this require that individual data files in the original record should be referencable and citable, it also means that the individual data elements in the files should retain referencability. This suggests that LTA Plans need to include a systematic discussion of the way in which such referencability will be maintained - and will require a discussion of the way in which new versions of the data will be able to be compared with the original versions. Furthermore, the Long Term Archive needs to ensure that data references and annotations retain their validity, so that interdisciplinary work that spans multiple data sets maintains traceability to the original data and does not have referencing that is destroyed or invalidated by data reformatting, reprocessing, or migration.

4. The Long Term Archival Plan needs to ensure that data is accessible to both NOAA and NASA scientific teams who are developing new instruments and new algorithms that extend the long term record of Earth observations.

In many cases, the most active period of data use occurs long after the original data sets have been produced, particularly when a new capability for extending a particular kind of measurement appears. New approaches to measuring a quantity often require comparing a historical data set with the new one, even after a gap of decades. Often, these comparisons require complete data sets of Level 1 and Level 2 data in the location where the new instrument teams do their work. It is important to incorporate the data flows that these activities require into the LTA design. For example, this may require LTA design for moving large data volumes quickly - so that five years of Level 2 data could be delivered in a few months. It may also require duplication of data sets at disciplinary data centers outside of the LTA to avoid long delays in validating new data products. The LTA Plan needs to incorporate a careful discussion of tradeoffs in speed of data delivery versus storage costs for data set replication.

[UWG Home Page](#)

